

## ЭЛЕКТРИЧЕСКОЕ ЭХО на последней миле

В.Гордиенко, д.т.н., декан МТУСИ  
С.Шаврин, д.т.н., профессор МТУСИ

Электрическое эхо в совокупности с другими эффектами, вызванными его некорректным подавлением, в современных телекоммуникационных сетях стало едва ли не доминирующим фактором мешающего воздействия на качество телефонной передачи. Как снизить это воздействие?

### НЕИЗБЕЖНАЯ ПОМЕХА

Эффект электрического эха обусловлен проникновением части энергии задержанного речевого сигнала из прямого направления передачи канала связи в обратное. И в не столь далеком прошлом нарушение эхосигналами естественности ведения разговоров рассматривалось как их единственный мешающий фактор, а степень этого воздействия количественно оценивалась минимальным значением затухания эхосигналов, обеспечивающим приемлемое качество телефонной передачи, – как функции времени их распространения. Вместо времени распространения часто указывалась протяженность соединения, поскольку основная задержка в сигнал вносилась линейным трактом систем передачи. Для борьбы с мешающим воздействием эффекта электрического эха использовался заграждающий принцип, предполагающий внесение значительного затухания в обратное направление передачи при передаче речи в прямом. Бытовало также мнение, что радикальное решение проблемы эффекта электрического эха на сетях связи – электрически или логически полностью четырехпроводное соединение. В целом же трафик по каналам, требующим защиты от мешающего воздействия эффекта электрического эха, составлял незначительную долю общего телефонного трафика, и проблема защиты не ставилась во главу угла.

Между тем в восприятии эхосигналов наблюдается ряд парадоксов.

### ПАРАДОКСЫ ВОСПРИЯТИЯ

Главный парадокс – факт качественного изменения характера зависимости восприятия эхосигналов от времени их задержки. При малых значениях времени задержки эхосигналы интерпретируются как местный эффект и оказывают положительное влияние на общее качество телефонной передачи. Экспериментально установлено даже оптимальное (в смысле наиболее высокой абонентской оценки) значение затухания эхосигналов местного эффекта – оно равно 19 дБ [1].

При том же значении затухания в 19 дБ те же самые эхосигналы, задержанные на 50 мс, воспринимаются говорящим абонентом как помеха, отрицательное воздействие которой на качество телефонной передачи требует обязательного использования специальных эхоподавляющих устройств (ЭПУ).

Если же время задержки превышает 200 мс, то мешающее воздействие эхосигналов на говорящего абонента оценивается эквивалентно мешающему воздействию обычного гауссовского шума примерно в 25 раз большей мощности [2].

Еще одним парадоксом восприятия эхосигналов является тот факт, что говорящий абонент не узнает в них собственный голос. И даже когда абонент понимает, что это эхо, тем не менее мешающее воздействие эхосигналов на абонента сохраняется. Эти и другие феномены восприятия эхосигналов в телефонных каналах подробно рассмотрены в [2].

Последняя ревизия зависимостей, определяющих допустимые характеристики эхосигнала [3], ввела в обращение эквивалент затухания:

$$L_c = \frac{1}{m} C - \frac{10}{m} \lg \left[ \int_{300}^{3400} \frac{10^{-A_e(f) \cdot m/10}}{f} df \right],$$

где  $C = 3,9$  дБ,

$$m = \begin{cases} 0,2 & \text{для речи} \\ 0,5 & \text{для эха слушающего,} \\ 1,0 & \text{для эха говорящего} \end{cases} \quad (1)$$

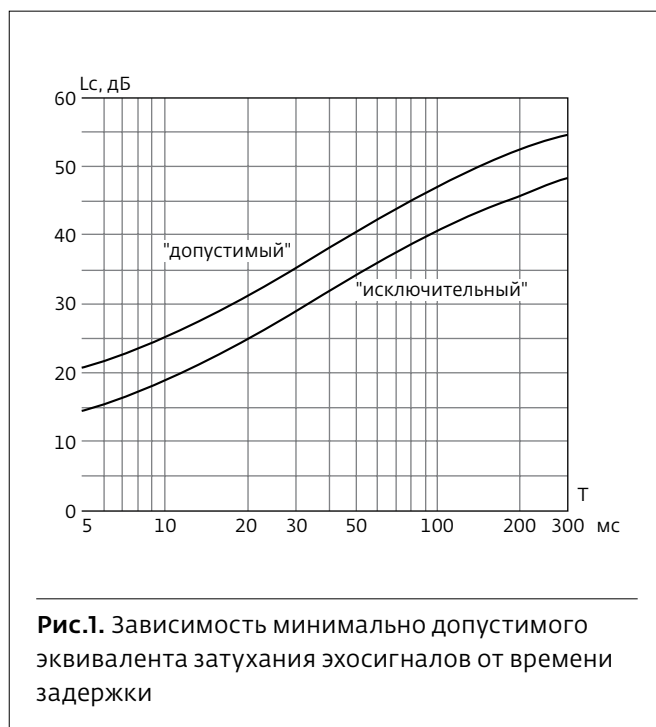
а ход зависимости этого эквивалента затухания от времени задержки эхосигналов, обеспечивающего приемлемую степень мешающего воздействия эффекта электрического эха, приобрел вид, представленный на рис.1.

Параллельно проводились исследования по оценке степени мешающего воздействия эффекта электрического эха по различным субъективным критериям с использованием метода мнений (MOS), заметности, затруднений и др. Многочисленные результаты проведенных исследований легли в основу так называемой E-модели [4], в рамках которой МСЭ ввел в обращение понятие показателя R качества телефонной передачи с упором на ее диалоговый характер (conversational quality), учитывающего совместное воздействие на качество основной группы мешающих факторов, в том числе эхосигналов. Кроме того, были разработаны методики и программные средства измерения качества телефонной передачи, учитывающие субъективный характер оценки PSQM (Perceptual speech quality measurement) и PESQ (Perceptual evaluation of speech quality) [5, 6, 7, 8].

### НОВЫЕ ПРОБЛЕМЫ И РЕШЕНИЯ

Цифровизация связи дала новый мощный импульс борьбе с мешающим воздействием эффекта электрического эха, превратив ее в системную задачу. Пакетный принцип передачи, распределенная коммутация, сжатие речи в стремлении обеспечения минимальной избыточности, – все эти факторы приводят к повышению времени распространения от источника речевого сигнала до потребителя и, как следствие, – к повышению степени мешающего воздействия эффекта электрического эха.

Помимо этого, IP-технологии принесли в телефонию целый ряд специфических явлений, ранее нехарактерных для подобного рода приложений. К таким эффектам прежде всего относится пропадание пакетов, усугубленное вызванным им же нарушением адекватности восстановления

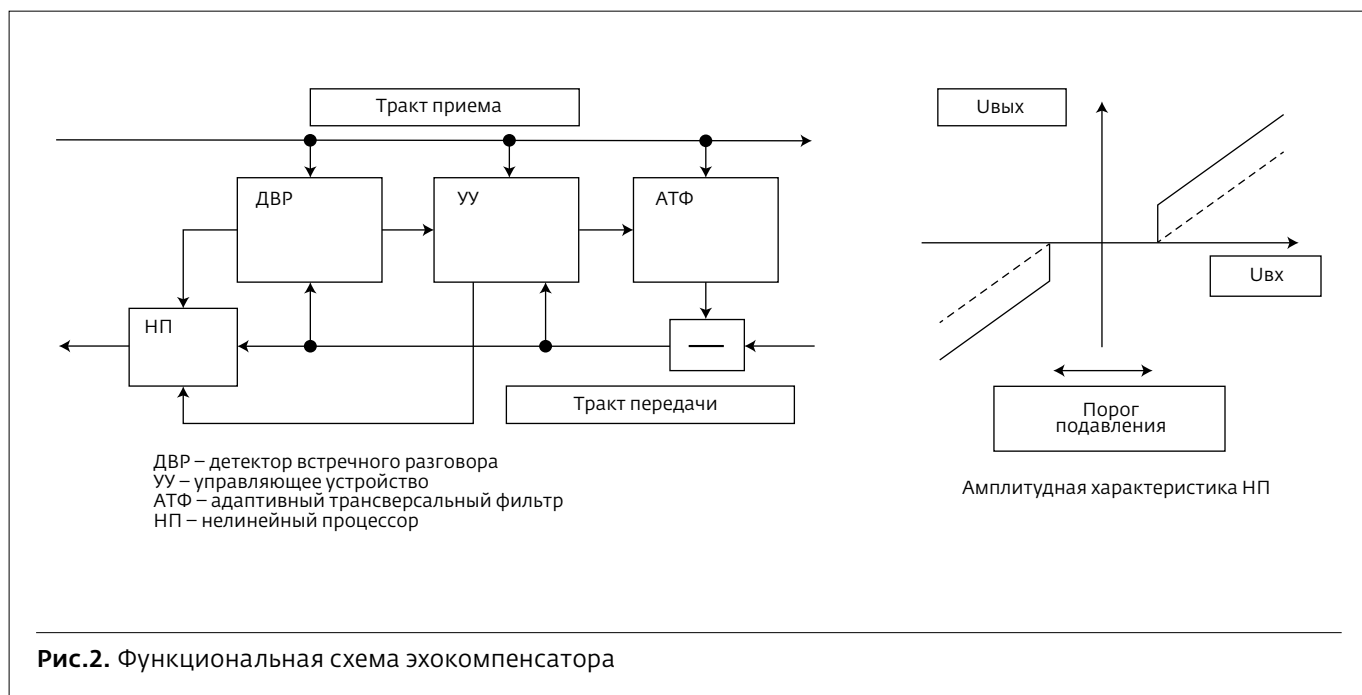


**Рис.1.** Зависимость минимально допустимого эквивалента затухания эхосигналов от времени задержки

сжатых речевых сигналов декодерами. В некоторых случаях последствие пропадания одиночного пакета может быть весьма значительным и достигать полусекунды, причем субъективное восприятие того сигнала, который попадает к абоненту, по мешающему воздействию превышает обычное сгруппированное размножение ошибок передачи [9].

Развитие цифровых методов обработки сигналов и цифровой элементной базы открыло возможность реализации альтернативного заграждающему компенсационного принципа подавления эхосигналов, основанного на формировании копии эхосигнала и ее вычитании из смеси эхосигнала и полезного сигнала обратного направления передачи.

Функциональная схема эхокомпенсатора (ЭК) представлена на рис.2; характеристики ЭК регламентируются рекомендациями МСЭ-T G.165 [10] и G.168 [11]. Исполнительный элемент ЭК, реализующий функцию формирования копии эхосигнала, строится на основе адаптивного цифрового фильтра с конечной импульсной характеристикой – трансверсального фильтра (АТФ). Эта тенденция сохранилась до настоящего времени с первых образцов ЭК, невзирая на некоторые теоретические разработки по использованию в этой функции фильтров с бесконечной импульсной характеристикой, лестничных и некоторых других видов фильтров.



**Рис.2.** Функциональная схема эхокомпенсатора

Формирование копии эхосигнала реализуется в соответствии с некоторым функциональным адаптивным алгоритмом, разработка или выбор которого относится к компетенции разработчика оборудования. Сходимость функционального алгоритма в тех или иных условиях эксплуатации определяет скорость и степень подавления эхосигнала, и в значительной степени общее качество телефонной передачи.

Потенциально ЭК может обеспечить весьма высокое общее качество телефонной передачи (даже более высокое, чем при полностью четырехпроводном соединении за счет сохранения для абонентов комфортного уровня местного эффекта). Однако в реальных условиях эксплуатации ЭК оказывается подверженным воздействию целой гаммы факторов, дестабилизирующих нормальную работу собственно компенсационного принципа подавления эхосигналов. К таким факторам прежде всего относятся: нелинейность эхотракта; шум в тракте передачи; наличие в эхотракте систем сжатия речи, использующих предиктивные механизмы; сдвиг спектра эхосигнала по отношению к исходному сигналу, вызванный погрешностью восстановления частоты в двухполосных двухпроводных системах передачи по соединительным линиям местных сетей; параметрические явления в эхотракте; низкое затухание эхосигналов и др.

В отличие от ЭК, используемых для дуплексной передачи данных по телефонным каналам,

адаптация голосовых ЭК осуществляется не по специальному тест-сигналу в специально отведенном отрезке времени, а непрерывно по речевому сигналу, не обладающему приемлемой степенью стационарности.

Результат неполной сходимости ЭК – наличие остаточного эхосигнала, заметного для абонентов. В системах связи, использующих сжатие речевого сигнала, остаточный эхосигнал (помимо своего собственного мешающего воздействия) генерирует еще один негативный эффект – "металлический" звук, вызванный нарушением адекватности преобразования суммы двух разных речевых сигналов вокодными алгоритмами [12].

## СПЕЦИФИКА VoIP

В терминологии сетевых приложений передача речевой информации по протоколам IP получила название VoIP (Voice over IP). Для передачи пакетов обычно используется протокол RTP (Real-Time Transport Protocol) [13, 14], как правило поверх UDP (редко TCP); механизм реагирования на изменения в сети поддерживается с помощью протокола RTCP (RTP Control Protocol). По умолчанию длина пакета устанавливается в расчете на длительность отрезка речевого сигнала, равную 20 мс. Метки времени, передаваемые в составе протокола RTP, обычно используются для правильного восстановления речевого сигнала из принятых пакетов. Управление процессом установления соединений

обычно осуществляется по протоколу SIP (Session Initiation Protocol) [15].

Время распространения речевого сигнала при использовании технологий VoIP может быть определено по методике, представленной в [16]. Следует, однако, иметь в виду, что основную долю времени распространения часто определяет задержка сигнала в буфере сбора (джиттера) пакетов. Этот буфер включен в состав каждого, в том числе индивидуального, шлюза и предназначен для восстановления правильного порядка следования пакетов, распространяющихся по разным маршрутам. Емкость буфера должна устанавливаться в соответствии с ожидаемой предельной задержкой пакета в сети, поскольку все пакеты, пришедшие позже, игнорируются. Значение емкости может устанавливаться вручную, однако многие виды шлюзов поддерживают функцию адаптивного управления емкостью в соответствии с реальной статистикой приема пакетов.

Ограниченная пропускная способность сетей наряду со стремлением к экономии ресурсов заставляет как операторов IP-сетей, обслуживающих абонентов сети общего пользования, так и абонентов, выходящих в сеть с индивидуальных шлюзов, повышать емкость буфера с целью минимизации потерь передаваемых пакетов. Поскольку в большинстве выпускаемых в настоящее время шлюзов емкость буфера соответствует задержке сигнала в 300 мс, это значение, как правило, устанавливается дополнительно к задержке, вносимой остальным оборудованием. Таким образом, общая задержка сигнала в сетях IP-телефонии вполне соизмерима с задержкой в спутниковых каналах связи и требует использования эхоподавляющих устройств.

Функция подавления эхосигналов в IP-соединениях осуществляется ЭПУ, встроенным в оборудование шлюзов. Как правило, встроенные ЭПУ выполнены в соответствии с [11]. Практика подавления эхосигналов в каналах IP-телефонии демонстрирует далеко не всегда достаточную эффективность работы встроенных ЭПУ.

При этом пропадание пакетов (до 5%), использование функции детектирования голосовой активности, сжатие речи и другие факторы нивелируют разницу в качестве телефонной передачи, достижимом при использовании заграждающего и компенсационного механизмов. Сложившееся положение заставляет многих операторов и потребителей услуг IP-телефонии использовать полнофункциональные ЭПУ заграждающего типа, обеспечивающее устойчивое двустороннее подавление

эхосигналов – как для своих абонентов, так и для их собеседников.

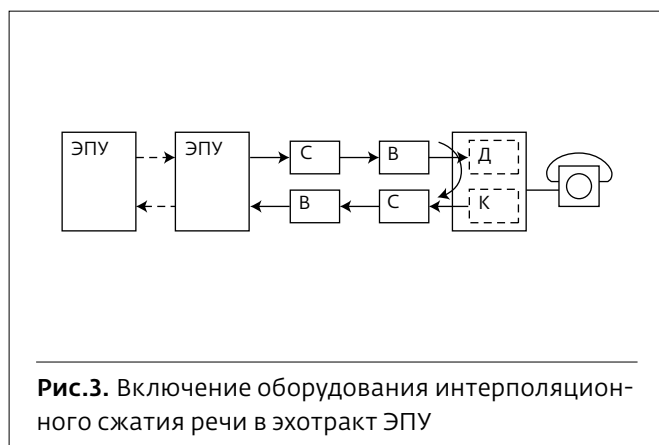
Главная особенность подавления эхосигналов в двустороннем режиме в соединениях IP-телефонии – существенное влияние изменений емкости буфера сбора (джиттера) пакетов в процессе разговора на эффективность работы ЭПУ. Эффективное двустороннее подавление эхосигналов в таких соединениях может быть обеспечено только при отключенной функции адаптивной настройки емкости буфера.

Неподавленные эхосигналы в каналах IP-телефонии будут оказывать отрицательное влияние как на эффективность (пропускную способность) сетей (аналогично влиянию на эффективность статистических систем передачи), так и на качество телефонной передачи при использовании сжатия речи. Для минимизации этого влияния встроенные ЭПУ шлюзов не следует отключать при организации двустороннего заграждающего подавления эхосигналов. Такой принцип подавления может считаться оптимальным для каналов IP-телефонии на современном этапе развития ЕСЭ РФ.

В соответствии с действующими правилами построения ЕСЭ РФ и возможностями используемых на сети систем сигнализации, реально формирование соединений, в которых оборудование сжатия речи окажется включенным в эхотракт со стороны станционных портов ЭПУ, как показано на рис.3 (обозначения: С – сжиматель речевого сигнала; В – восстановитель речевого сигнала; К – кодер системы передачи; Д – декодер системы передачи).

Такой вид соединений далеко не всегда сможет обеспечить приемлемое качество телефонной передачи. Так, компенсационный механизм подавления эхосигналов, прошедших процедуру сжатия и восстановления интерполяционными механизмами, не сможет обеспечить требуемую степень подавления. В соответствии с результатами исследований, представленных в [16], предельная степень подавления эхосигналов в этом случае составит величину порядка 12 дБ. Заграждающий механизм, обладая высокой чувствительностью к воздействию импульсной помехи, обеспечит высокое качество телефонной передачи только при очень низком (ниже  $10^{-8}$ ) коэффициенте ошибок передачи сжатого речевого сигнала.

Изложенные соображения дают основания считать подобные соединения неприемлемыми в качестве систематического явления на сети и не использовать их в коммерческой эксплуатации



без крайней необходимости. На направлениях, где отказ от подобных соединений проблематичен, лучше использовать заграждающий принцип для подавления преобразованных эхосигналов. При возможности ЭПУ следует разместить ближе к уплотненному участку.

## БОРЬБА ПРОДОЛЖАЕТСЯ

Современные средства сжатия речи представляют собой сложные системы, функционирующие по принципу моделирования речеобразующего аппарата человека. Получившие наиболее широкое распространение на телекоммуникационных сетях алгоритмы сжатия RPE-LTP (в сотовой подвижной связи) и CS-ACELP (в наземных, в том числе IP) функционируют по принципу "анализ через синтез", обеспечивая с помощью системы фильтров и генератора реализацию сигнала, максимально близкую к сжимаемой.

Оптимизация процессов анализа и синтеза речи в системах, использующих основные рассмотренные алгоритмы, ориентированная на соответствие свойств кодируемого сигнала свойствам одиночного голоса (речевого сигнала), будет в значительной степени нарушена при наличии в составе кодируемого процесса эхосигнала, представляющего собой другой голос соизмеримой мощности.

Синтез декодером системы сжатия двухголосного сообщения, представляющего собой смесь речевого сигнала одного абонента с эхосигналом другого абонента, будет вызывать значительные взаимные влияния между полезным сигналом и эхосигналом и нарушать адекватность восстановления речи при передаче по цепи кодер - декодер. Искажения, вносимые в передаваемый сигнал в процессе кодирования и декодирования системами сжатия, вызванные влиянием эхосигналов, будут усугубляться и растягиваться во времени механизмами

последствия, заложенными в процедурах сжатия, а также влиянием ошибок передачи.

Прецизионные свойства используемых при моделировании алгоритмов и механизмов ориентированы на моделирование одиночного голоса, включая, например, фазу генератора основного тона в текущей реализации. Наложение эхосигнала - голоса другого абонента - на моделируемую реализацию нарушает адекватность процессов преобразования, приводя к появлению в восстановленном речевом сигнале погрешности, которую можно интерпретировать как дополнительный случайный процесс. Этот процесс, не являющийся линейной или простой нелинейной комбинацией исходных сигналов, имеет характерное "металлическое" звучание и остаточную разборчивость, его пикфактор имеет порядок 28...33 дБ, а пиковые значения могут превышать уровень основного речевого сигнала даже при высокой (~27 дБ) степени подавления эхосигнала [12].

Высокая чувствительность интерполяционных алгоритмов сжатия речи к влиянию неподавленных эхосигналов может вызвать нарушение адекватности преобразования и при одностороннем разговоре, если на эхосигнал будет накладываться разговор или шум помещения ближнего по отношению к кодеру абонента.

## ЛИТЕРАТУРА

1. **Вемян Г.В.** Передача речи по сетям электросвязи. - М.: Радио и связь, 1985.
2. **Шаврин С.С.** Эффект электрического эха в телекоммуникационных системах. - М.: Инсвязыздат, 2010. 178 с.
3. ITU-T. Recommendation G.131. Talker Echo And Its Control. ITU, 2003.
4. ITU-T. Recommendation G.107. The E-model, A Computational Model For Use In Transmission Planning. ITU, 1998.
5. ITU-T Recommendation P.861. Objective quality measurement of telephone-band (300-3400 Hz) speech codecs. ITU, 2008.
6. ITU-T Recommendation P.862. Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs. ITU, 2001.
7. ITU-T. Recommendation P.800. Methods for subjective determination of transmission quality. ITU, 1996.
8. ITU-T. Recommendation P.831. Subjective performance evaluation of network echo cancellers. ITU, 1998.



9. **Шаврин С.С.** Эффект электрического эха как системная проблема для операторов мобильной связи // Вестник связи. 2008. № 4. С. 108-112.
  10. ITU-T. Recommendation G.165. Echo cancellers. ITU, 1993.
  11. ITU-T. Recommendation G.168. Digital network echo cancellers. ITU, 2009.
  12. **Гордиенко В.Н., Федоткин Н.Н., Шаврин С.С.** О некоторых результатах исследования влияния эхосигналов на процесс передачи речи при использовании методов сжатия // Тез. докл. науч. конф. профессорско-преподавательско-го, научного и инженерно-технического состава. – М.: МТУСИ, 2005. кн. 1, с. 172.
  13. RFC3550 RTP: A Transport Protocol for Real – Time Applications. July 2003.
  14. RFC3551 RTP Profile for Audio and Video Conferences with Minimal Control. July 2003.
  15. RFC3261 SIP: Session Initiation Protocol. June 2002.
  16. **Шаврин С.С.** Развитие теории и техники подавления эффекта электрического эха в телекоммуникациях // Дисс. на соиск. уч. степени докт. техн. наук. Москва, 2009.
-